***HW3***

1. The **K**-**means** clustering algorithm is **sensitive to outliers**, because a **mean** is easily influenced by **extreme values**. ... **Mean** is greatly influenced by the **outlier** and thus cannot represent the correct cluster center, while medoid is robust to the **outlier** and correctly represents the cluster center
2. Will define:

Will differentiate with respect to

And get:

Will substitute and will get

Therefor it’s a extremum point will show it’s a minimum rather than a maximum

Will differentiate again: so the 2nd derivative is always positive and thus the function is convex and have only minimum point.

1. Will define therefor

Will differentiate with respect to

And get:

Will substitute and will get

Will open:

We know that exactly half of the x’s bigger than the median and exactly half of them smaller therefore will get exactly half of the numerator are +1 and exactly half are -1

And all of the denominator are +1 there for will get something like that:

Will differentiate again: so the 2nd derivative is always positive(or 0) and thus the function is convex and have only minimum point.

1. We have 2 SVM with linear karnel A and D . A as samples within the margins and thus have more soft margins and thus smaller penalty term C.

A-1

D-2

C and F are polynomial karnel, and its clear F is with higher degree.

F-4

C-3

B and E are both rbf, and B boundary is tighter so its with bigger ,which is correlated with tighter boundary.

B-6

E-5

1. The scientific term is generalization
2. 2p could imply about overfitting (lack of generalization) model if its 2 big.

And 2ln(L) could imply about underfitting (lack of generalization) if its to small.

Therefor a good AIC is if there is a balance between the 2 parameters.

1. If the balance is violated there could be under/overfitting.
2. AIC is a criterion that help you asses goodness of a model. Lower values of the index indicate the preferred model, that is, the one with the fewest parameters that still provides an adequate fit to the data.